FINAL DRAFT

REPORT OUT

TO THE

MODELING AND SIMULATION WORKING GROUP

Prepared by

Modeling and Simulation Benefits Task Force
Defense Modeling and Simulation Office

December 5, 1995

EXECUTIVE SUMMARY

The Modeling and Simulation Working Group (MSWG) authorized the formation of a Modeling and Simulation Benefits Task Force (MSBTF) to capture documented reports of quantifiable benefits of modeling and simulation. This effort is an initial step in fulfilling sub-objective (6-1) of the Department of Defense Modeling and Simulation Master Plan, to quantify the impact of modeling and simulation, and will be input directly to the longer-range Impact Assessment. The MSBTF first met February 23, 1995, continued meeting monthly, and stood down on September 21, 1995. This report describes the task force's efforts and findings.

During the Task Force's "data capture" period, responses were received from two computerized Requests for Information (RFI), task force members collected inputs from their respective Services and Agencies, and formal studies were collected from Federally Funded Research and Development Centers (FFRDCs) and Components. The RFI conducted for the Task Force produced a dozen responses. A separate RFI conducted for the test and evaluation community produced fifteen responses. Overall, Task Force members collected over 90 contributions.

While no formal assessment could be accomplished based on the relatively small amount of information gained in such a short period of time, an informal meta-analysis—an analysis of other organizations' analyses—is included in this report. Meta-analysis has two serious shortcomings: it is based on information often a year or more old, and it severely underreports negative results. The Task Force provided some old and some very new inputs. FFRDC inputs, while recently published, were sometimes based on data collected in prior years.

Findings

The applications of M&S to acquisition are many. Twenty-one case studies of Target Interaction, Lethality and Vulnerability showed a 30 to 1 return on investment in M&S support for milestone decisions and the Cost and Operational Effectiveness Analysis (COEA) process. The Army Missile Systems Command reported over $320 million in cost avoidance and savings from 10 case studies. Eight additional case studies were provided from the Virtual Proving Ground. Two similar events were conducted for Apache Longbow Force Development Test and Experimentation, one using extensive simulation and the other using physical equipment. The simulation-supported event executed twice as many trials, with fewer personnel, in less time, at lower risk to personnel, for $700,000 versus $4 million.

Training applications of M&S were common and positive. Individual skills training, including both cognitive and psychomotor skills, were well reported. Cognitive skills trainers, typically computer-aided instruction, paid for themselves in five years or less. Psychomotor skills trainers, e.g., flight simulators, driver trainers, conduct of fire trainers, and maintenance trainers, were all shown to be cost effective when properly mixed with training on the real equipment. Analysts have well-established theories and experimental methods for conducting analysis at this level. The same is not true for training units, particularly high echelon units. The high cost of a Joint or Combined exercise precludes the repeated, controlled experiments necessary to gather meaningful data. However, multi-million dollar savings are reported when comparing computer-assisted command post exercises to field training exercises.

Although M&S is used extensively in analysis, few reports were submitted. No reports were received claiming cost savings. The training and acquisition functional areas are more likely to gather quantitative data, specifically cost data. Information provided from members of the analytic community suggests that they tend to measure the effect of their analysis on the decisionmaking process rather than the effect of M&S on their analysis. The effect of M&S on analysis, while real, is problematic to quantify.

Across functional areas, measures of effectiveness (MOEs) are not universally agreed upon. Consistency within functional areas is somewhat better. Analytic frameworks, including MOEs, need to be developed and applied consistently throughout DoD. Frameworks should be unique to each functional area, or perhaps even functional sub-areas. Analytic frameworks will be recommended best practices.

Finally, no formal reporting mechanism exists for gathering information, nor does there exist a process for objectively assessing gathered information. If the effects of M&S are to be collected, assessed, and disseminated, then reporting pipelines should be established from the M&S developer and user communities through the Components to the Defense Modeling and Simulation Office.

Table of Contents

List of Tables

Acknowledgments

The efforts and contributions of the individual members of the Modeling and Simulation Benefits Task Force, under the direction of LtCol David A. Bartlett, USMC, are gratefully acknowledged. Authorship of this document belongs to the following Task Force members.

Dr. D. Robert Worley	Institute for Defense Analyses
Dr. Henry K. Simpson	Defense Manpower Data Center
Mr. Matt Aylward	MITRE Corporation
Ms. Michelle Bailey	Naval Air Warfare Center Weapons Division
Major Dean Fish	Marine Corps Modeling and Simulation Management Office

INTRODUCTION AND SUMMARY

Tasking from DoD M&S Master Plan

The Department of Defense Modeling and Simulation Master Plan contains six broad objectives. Objective 6, "Share the benefits of M&S," the basis for the Task Force, has three sub-objectives:

- 6-1: "Quantify the impacts of M&S"
- 6-2: "Educate potential M&S users (DoD, Congress, other government agencies, industry)"
- 6-3: "Support bi-directional technology transfer"

Sub-objective 6-1 further states that to quantify the impacts of M&S, the following actions must be accomplished.

- Collect and analyze data from ongoing efforts, planned experiments, and demonstrations to assess the impacts of M&S.
- Develop quantitative measures of the benefits of M&S (e.g., readiness impact, cost savings, and effectiveness) to (1) allow assessment of the utility of M&S and (2) support investment decisions.
- Establish the DoD-wide impact of M&S.

Method

The Defense Modeling and Simulation Office has initiated three activities in response to tasking: (1) The Institute for Defense Analyses has collected information on 48 projects and 17 activities sponsored by DMSO in fiscal years 1992 through 1994, (2) the Task Force has conducted an initial data capture from the Components, and (3) a sustained, long-range Impact Assessment has begun employing the previous efforts as input.

Report Overview

The second chapter of this report provides the Task Force's composition and schedule as well as its major findings and conclusions.

Quantified impacts for M&S applications in acquisition, training, and analysis, are summarized in the remaining chapters.

Measures of Effectiveness (MOEs) are not universally agreed upon across functional areas. Consistency within functional areas is somewhat better. Appendices A and B contain two Task Force members' surveys of quantitative and qualitative MOEs. Appendix B emphasizes methodologies.

Finally, an annotated bibliography covering the full range of material collected by the Task Force is provided with full bibliographic references, when possible.

Findings

No formal reporting mechanism exists for gathering information, nor does there exist a process for objectively assessing gathered information.

Where reported, quantified impacts are significant. No examples of marginal or negative impacts were found. Note: this is a typical shortcoming of conducting an analysis of other organizations' analyses.

The quality and methods of analysis are not uniform. Two members of the Task Force surveyed their own areas to identify measures of effectiveness and methods of analysis employed. Appendices A and B contain the results of their surveys.

The training and acquisition functional areas are more likely to gather quantitative data, specifically cost data. Information provided from members of the analytic community suggests measurement of the impact of their analysis on the decisionmaking process, rather than the impact of M&S on their analysis. Quantifying the M&S impact on analysis, while real, is problematic. Since the Army has devoted significant resources to M&S in the past few years, it appears more likely to measure and report the impact of M&S than the other Services.

Conclusions

To measure the impact of the many applications of M&S in the Department of Defense, objectives must be first stated in quantifiable terms. Only then can our progress toward attaining those objectives be assessed.

Analytic frameworks, including measures of effectiveness, need to be developed and applied consistently throughout DoD. Frameworks should be unique to each functional area, and perhaps even functional sub-areas. Analytic frameworks should developed within the functional communities and serve as recommended best practices.

Reporting pipelines should be established from the M&S developer and user communities feeding into DMSO for collection, assessment, and dissemination.

Task Force Timeline

Table 1: M&S Benefits Task Force Timeline

TF meeting (kickoff)	February 23, 1995
Interim results distributed	March 10, 1995
Meeting, submit and discuss major points	March 23, 1995
Meeting, agree on major report points	April 20, 1995
Draft report distributed for TF review	May 15, 1995
Meeting (final), discuss draft report	May 25, 1995
Final report completed	September 21, 1995

Task Force Membership

The membership of the MSBTF was drawn from the Services, OSD offices, Federally Funded Research and Development Centers (FFRDCs), and private industry. LtCol David A. Bartlett, DMSO, chaired the task force. Composition of the task force is shown in Table 2 below.

Table 2: M&S Benefits Task Force Members

Organization Represented	Member
Mitre, Quantico	Mr. Matt Aylward
NAWCWD, China Lake	Ms. Michelle Bailey
DMSO	LtCol David A. Bartlett, USMC
IST/UCF	Mr. John Bishop
DAMO-TRS, DCOPS	Colonel Guy M. Bourn, USA
AEDC/DOF	Mr. Tom Brown
OUSD(A&T) DTSEE/MSEE	Mr. Albert R. Burge
DATAMAT Systems Research	Mr. Bob Curtis
Contraves, Inc.	Mr. Michael L. Darby
MCMSMO	Major Dean E. Fish, USMC
Loral Federal Systems	Mr. Trevor Huth
OUSD (P&R)	Mr. Don Johnson
ARI	Dr. Frank Moses
OUSD(A&T), DTSEE	Mr. Fred A. Meyers
IDA	Dr. Jesse Orlanksy
DISA/D84	Mr. Ron Prishivalko
DMDC	Dr. Henry K. Simpson
AEDC/DOF	Mr. J. Stephen Schroeder
IST/UCF	Mr. Ernie Smart
NAWCWD, China Lake	Dr. Robert D. Smith
IDA	Dr. D. Robert Worley

ACQUISITION APPLICATIONS OF M&S

Target Interaction, Lethality, and Vulnerability (TILV)

TILV refers to the science of understanding the mechanisms by which a warhead or other damage mechanism can defeat a target [TILV 1995]. The TILV area addresses the tools, methods, databases, and supporting techniques needed to assess the lethality and vulnerability of all weapon systems, including aspects of design, effectiveness, and survivability. TILV tools provide essential inputs to milestone decisions and the COEA process. TILV analyses play a crucial role throughout the development cycle of all major military systems. Table 3 depicts the Return on Investment (ROI) of 20 systems. ***The typical ROI was between $20 and $30 returned for each $1 invested***.

Table 3: TILV Return on Investment

Program	Tool	Type	Total Invest ($M)	Direct Savings ($M)	ROI	Program Result
AMRAAM	SHAZAM	End Game	6.5	250	38	Continued
MK Series Bomb Fragment Data		Arena Tests	.0825	.9	11	Continued
BLU-109		Lethality Testing	.0825	3	36	Continued
Air-to-Air Missile Analysis	ACEVAL/ AIMVAL	Lethality plus Engagement	20.0	75	4	Continued
Wide Area Anti-Armor Munition		Lethality Analysis	.75	30	40	Canceled
Hypervelocity Missile		Lethality Analysis	.5	10	20	Canceled
ISAS		Lethality Analysis	.75	40	53	Canceled
Kinetic Energy Penetrator (KEP)		Lethality Analysis	1.1	50	45	Canceled
JP 233 Runway Attack Munition		Lethality and Vulnerability Analysis	1.1	54	49	Canceled
Boosted Kinetic Energy Penetrator		Runway Vulnerability Models	2.75	130	47	Canceled
JAVELIN ATGM		Analytic Simulation	.62	14	23	Accepted
M2 Bradley FVS		Engineering Design	.88	30	34	Accepted
Abrams M1A2 Vulnerability Test		Damage Prediction	1.83	30	16	Less Test & Damage
Block 3, M1A2		Design Vulnerability	1.76	100	57	Terminated
Standard Missile SM-2 BLK IIIA	COVART WHDEVAL	Cost Reduction	2.25	47	21	Accepted
PHALANX CIWS		Performance Evaluation	8.12	125	15	Continued
PHALANX CIWS Upgrade		Product Upgrade	6.63	200	30	Accepted
AIM-7P Sea Sparrow	SCAN	Lethality Analysis, End Game	.7	16	23	Accepted
Phoenix Missile	SCAN	Lethality Analysis, End Game	2.23	70	31	Accepted
ECM vs. AMRAAM	SCAN	Lethality Analysis, End Game	.58	10.5	18	Evaluations Continue

TECOM Virtual Proving Grounds (VPG)

For modeling and simulation to be useful and valid, the applicable tools must be based on real data derived from testing—ground truth [TECOM 1995]. The Army's Test and Evaluation Command (TECOM) supports this concept with an approach known as Simulation and Modeling Anchored by Real Testing (SMART). TECOM is developing the VPG as a means for researchers and developers to assure that their models and simulations are based on real test data.

VPG is a network of models and simulations enabling interactive testing within a synthetic environment. A number of projects undertaken by TECOM use these models and simulations to determine the various effects on systems and to replicate actions without undertaking the time and expense of actual testing.

Table 4 represents a summary of selected TECOM systems that used VPG in conducting tests and evaluations with cost avoidance as a measure of effectiveness. Actual cost includes investment in simulation when appropriate and available.

Table 4: TECOM VPG Cost Avoidance

Project	Use	Simulation	Actual Cost ($M)	Cost Avoidance ($M)
Firing Impulse Simulator	Recoil loads and ballistic shock effects	Replicate actual firing without the use of ammunition for tanks and howitzers	6.9	23
M830E1 Fuse Testing	Evaluates tank vs. helicopter engagements	Virtual test range simulation using simulated helicopter engagements with actual manned tank	.26	1.5
Moving Target Simulator	Immersion of entire weapon system (air or ground) into moving visual target environment	Assess the ability of an M1A2 tank crew to track and simulate firing on images of simulated maneuvering targets		1.5 per year
Simulation/Test Acceptance Facility (STAF)	Test millimeter wave radar-guided missiles	Hardware-in-the-loop simulator providing test of a fully assembled "live" missile with multiple computer-based test scenarios		10.6 per year
Aerial Cable Range (ACR)	Test missile tracking of heat sources	Uses a 3-mile long suspended Kevlar cable that serves as path for captive vehicles	.7	13.8
Test Item Stimulators (TIS)	Non-radiating simulated digital message traffic to C3 systems	Test of Enhanced Position Location Reporting System (EPLRS)	4.7	2
Trajectory Sense and Destroy Armor Simulation (SADARM)	Model ballistic simulation for the SADARM projectile	Enables downrange auto-trackers to acquire and track incoming projectiles and transition quickly to acquire end-game data		12
Physical Simulation of Bridge Crossing	Bridge durability tests	Mix of physical and simulated bridge crossings	.325	.11

Army Missile Systems

The Army Missile Command (USAMICOM) Research, Development and Engineering Center (RDEC) uses modeling and simulation extensively in the development of Army Missile Systems [Jolly and Ward 1995]. During the course of numerous simulation projects, the benefits of hardware-in-the-loop (HWIL) simulations have translated into cost savings and avoidance for many weapon system development programs. Examples of **cost saving and avoidance exceeding $320M** are presented in Table 5.

Table 5: Army Missile Systems Cost Effects

Project	Application of HWIL and DIS Simulation	Save/Avoid ($M)
MLRS-TGSM	45% reduction in flight/drop test program	6
FOG-M/NLOS	HWIL simulation identified all hardware and software faults prior to flight tests resulting in reduction in flight test costs	15
LONGBOW	Successful Proof-of Principle and EMD flight test programs with prevention of at least 2 test failures and reduction of risk in several other cases	6.5
Classified Program	Viability of this development program possible only through HWIL simulation; estimated flight test cost savings	60
HAWK	Flight test cost savings on counter ECM and other system PIPs	80
STINGER	Flight test cost savings for benign, countermeasured, and untestable scenarios	> 90
ATACMS	Analysis of flight test anomaly possible only with HWIL simulation; rapid identification of anomaly source saved extensive investigation	0.5
JAVELIN	Performance assessment data for milestone 3 decision produced by simulations, avoiding several flight tests	5
Foreign Materiel Exploitation	ECM hardware/software/techniques evaluation and optimization against foreign threat missiles (Desert Storm payoff in identified saving of at least one aircraft and pilot)	> 25
FAADS-BSFV (DIS)	Evaluate options using real soldiers, without requiring costly development of prototype systems, and save substantially on field testing	32.1
	Total Cost Savings or Avoidance	**> 320.1**

Apache Longbow

An example of some of the benefits of using modeling and simulation in the Apache Longbow program are summarized Table 6 below [Swinsick 1995]. Phase I of Force Development Test and Experimentation (FDT&E) was based on manned simulation. Phase II employed approximately the **same test scenario and activities** but used live equipment. **Twice as many trials were conducted in Phase I than in Phase II, at less cost, with fewer personnel, in less time.** FDT&E allows the helicopter crews to train on the new equipment without the risk associated with flying real equipment. It allows development and practice of new tactics, techniques, and procedures. Those responsible for developing scenarios for Initial Operational Test and Evaluation (IOT&E) have the opportunity to structure the very expensive operational test to gain the most critical information.

Table 6: Apache Longbow FDT&E Effects

Apache Longbow Force Development Test and Experimentation (FDT&E)		
Resources	**Phase I Manned Simulation**	**Phase II Field Test**
Cost (O&M Army)	$.712M	$4.049M
Equipment	1 Simulator	4 AH-64D 2 UH-60 14 M1 Tanks 10 M3 Fighting Vehicles 2 2S6 20 + Air Defense Units 47 + Vehicles
Personnel (Government)	27	663
Mission Turn-Around Time	2 Hours	6 Hours
Data Reduction Time	4 Hours	80 Hours
Number of Trials	32	16
Test Period	4 Weeks	6 Weeks
Safety	No Risk	Moderate Risk

TRAINING APPLICATIONS OF M&S

Individual Skills Training

Individual training is supported most often by *stand-alone simulators*. These simulators range from simple devices (such as rifle marksmanship trainers) to more complex devices (such as maintenance simulators, tank gunnery simulators, and flight simulators). [Simpson et al. 1995] drew these general conclusions about the effectiveness and cost of such simulators:

> ...in aggregate, simulators provide significant beneficial transfer from simulator to aircraft at a median operating cost of about one-tenth of an aircraft...**Because of their scope, the body of studies probably provides the strongest case for the value of *any* type of simulation**.
>
> Students trained using maintenance simulators perform about as well as those trained with actual equipment, but simulators cost a fraction...of the equipment...where time to train was reported, training with simulators took...less time than with actual equipment.

Aviation

- The Army estimates that it substitutes simulation for **$68M of flight operations training in the active force and $55M in the Reserves each year**; the Navy considers simulation to be effective in initial training in unfamiliar aircraft, as is reflected in the ratio of simulator to actual aircraft training flights (40 to 77) in the fleet replacement training program for F/A-18 aircraft; the Air Force Air Mobility Command plans to **replace up to 50 percent of flight training hours with flight simulators** and other training devices for training air transport crews [Orlansky et al. 1994] [Department of the Army 1993].
- The operating cost of flight simulators is estimated to be between 5-20% of the cost of aircraft. Many studies have shown that skills learned in flight simulators can be performed successfully in aircraft; the use of simulators for training can reduce flight time [Orlansky & String 1977]. In a more recent study, **the median cost ratio of simulators to aircraft was estimated to be 8%** [Orlansky et al. 1984].
- A review of several studies showed that the operating costs of flight simulators are about 10% of actual equipment per hour trained or 33%, if acquisition cost is taken into account. The majority of tasks trained on simulators (59%) have significant positive transfer to flight performance [Angier et al. 1993].
- Bombing and air drop accuracy data indicate that additional simulator hours seem to have a greater positive effect than additional flying hours, and simulator hours cost at most a third as much. **Helicopter accident data indicate that both flying hours and simulator hours reduce accidents, but simulator hours do not increase exposure to risk** [Horowitz et al. 1992].

Small Arms

- Several studies relating to the use of simulation in lieu of live fire indicate that performance with **simulation is at least equal to live fire training, but that cost is lower**. Soldiers with MACS (Multipurpose Arcade Combat Simulator) training expended **less rounds** during live-fire qualifications and **fewer soldiers failed to qualify** as compared to those trained using traditional methods. Several studies with the Squad Engagement Training System (SETS) have shown positive transfer from SETS to live fire. Training with the Indoor Simulated Marksmanship Trainer (ISMT) has been demonstrated to benefit live-fire performance. The Precision Gunnery Training System (PGTS), an inexpensive trainer for TOW and Dragon missiles, whose rounds are prohibitively expensive ($11,500 and $19,145, respectively, per round), has been demonstrated to be cost-effective, and **also permits training that would otherwise cost several hundred million dollars per year** if actual missiles were used [Bailey & Hodak 1994] [Wilhoite 1993] [Eisley et al. 1990] [Schendel et al. 1984] [Berg et al. 1993b].

Maintenance

- A review of maintenance simulators found that they are as effective for training as actual equipment trainers when measured by student achievement in school. In the majority of cases examined, the cost to develop and fabricate one unit was less than 60% of actual equipment and the cost of fabricating a second unit was less than 20%. Acquisition and use of a maintenance simulator over a 15-year period costs 38% as much as actual equipment. In studies where time to train was reported, simulators took 25-50% less time than actual equipment [Orlansky & String 1981].

Collective Skills Training

Collective training focuses on tasks performed collectively by groups of individuals (e.g., crews, teams, units) who must work together and coordinate their activities. The size of a collective may vary greatly and hence collective training varies considerably in scale. It is supported most commonly by *live* or *virtual* simulation. Some *stand-alone simulators* train smaller personnel collectives (e.g., flight crews, tank crews). *Advanced distributed simulation*—a set of disparate models or simulations operating in a common synthetic environment perhaps composed of live, virtual, and constructive simulations—can also be used for collective training. A recent analysis by the General Accounting Office [GAO 1993] conducted for Congress cited as exemplary several simulations used by the Army for collective training: COFT (Conduct of Fire Trainer), used on tanks and Bradley Fighting Vehicles; MILES (Multiple Integrated Laser Engagement System), used to simulate direct fire weapons from rifles to tank and helicopter gunnery systems; and SIMNET, used to provide crew-, platoon-, and company-level training. A 1994 review of technologies supporting virtual simulation indicated that it is becoming increasingly powerful and cost-effective [Office of Technology Assessment 1994].

Crew/Team

- Evaluations of the UCOFT (Unit Conduct of Fire Trainer) have been positive. **Tank gunners trained with UCOFT fire their opening rounds about 25% faster than conventionally trained gunners**. Based on an analysis of a hypothetical force-on-force engagement, UCOFT-trained gunners would be expected to kill significantly more opposing tanks than conventionally-trained gunners [Operational Research and Analysis Establishment 1990]. [Boldovici et al. 1985] reviewed UCOFT tests and concluded that UCOFT provides improvements in gunner proficiency. Substantial gains were found in percents of targets acquired, engaged, hit, and killed for groups undergoing sustainment and transition training. Gains were attributed to improvements in acquisition time, engagement time, and first round hits, which in turn allowed time to scan, acquire, and engage available second and third targets. [Hughes et al. 1988] evaluated the training effectiveness of the UCOFT empirically with 369 tank commander-gunner pairs and found that UCOFT training increased progress and improved performance on Table VIII and was well accepted by users.
- In tank gunnery, the introduction of the Conduct of Fire Trainer **reduced the annual expenditure of ammunition from 134 to 100 rounds per tank and improved marksmanship**. This resulted in an **annual cost avoidance of approximately $29M**. The new Tank Weapons Gunnery Simulation System is expected to reduce the annual consumption to 78 rounds, **for an additional saving of $21M to $50M each year** [Orlansky et al. 1994] [Department of the Army 1993] [Morrison et al. 1991] [Turnage & Bliss 1990].

Multiship Air Combat

- In evaluations of developmental DIS systems designed to support multiship air combat training in a combat engagement simulation environment, participating pilots and air weapons controllers indicated that simulation enhanced their combat readiness and was more beneficial in some areas than traditional unit training [Bell & Crane 1992] [Houck et al. 1991].
- In evaluations of a SIMNET-compatible air combat simulator, pilots received training and then rated their interest in receiving additional training on each of 30 tasks. Tasks with the highest rated interest can usually be practiced only in large exercises or cannot be practiced except in simulation. It was concluded that multiplayer simulator based training is a valuable training medium for increasing wartime readiness, especially for less experienced pilots [Crane & Berger 1993].

Tactical Ground Combat

- During the Persian Gulf War, at the battle of "73 Easting", U.S. troops destroyed an opposing force three times their size while fighting in an area the Iraqis had previously used for training exercises. Leaders of the U.S. force attributed their success to training they had received with live simulation, virtual simulation, and stand-alone crew training simulators [Orlansky 1993].
- The Army Science Board has estimated that simulators would enable it to reduce aviation and vehicle OPTEMPO (Operating Tempo) and training ammunition by 15-20% while maintaining the same or better level of unit performance [Army Science Board 1989].
- A series of SIMNET tests and evaluations have demonstrated its value for collective training. [Schwab & Gound 1986] evaluated SIMNET's capability to support platoon level command and control exercises to train individual and collective tasks. Three of the four platoons in each group, SIMNET and baseline, improved their performance between the first and second set of STXs (Situational Training Exercises). The SIMNET group **improved their average group score by 13% while the baseline group improved its score by 6%.** Findings of [Kraemer & Bessemer 1987] suggest that SIMNET training helped units develop and improve their fire control distribution plans and helped unit leaders develop the command, control, and communication skills to effectively execute those plans during platoon battle runs. [Brown et al. 1988] found that SIMNET training increased field exercise platoon performance, command and control, and leadership skills, and adequately portrayed vehicle and battlefield sounds. SIMNET also improved performance of command and control, platoon movement, leadership, and fire distribution during the company team ARTEP. [Burnside 1990] found that **35% of Army Training and Evaluation Program (ARTEP) Mission Training Plan (MTP) tasks can be trained with SIMNET**. [Bessemer 1991] found positive transfer of tactical training from SIMNET to field training. Analysis of an effectiveness comparison between SIMNET and home-station field training indicates that SIMNET is extremely effective in increasing performance for SIMNET-trainable tasks relative to field training. Tradeoff analyses showed that investment in SIMNET-like facilities could be repaid by an **8% to 14% decrease in OPTEMPO** [Angier et al. 1993].
- An analysis of the training capabilities and cost-effectiveness of the CCTT concluded that it has the potential to train tasks relating to command, control, and communication; maneuver and navigation; and teamwork and leadership. When fielded, CCTT would be cost-effective and **its life cycle costs would be paid back fully during its service life** [Noble & Johnson 1991].

Multi-Service and Joint Training

- ARI has successfully demonstrated the use of virtual simulation for multi-Service Close Air Support training and is currently expanding its demonstration platform to include the Joint fire support mission [ARI 1995] [Hawley et al. in press].
- Virtual simulation has the potential to enable Joint and inter-Service training in mission areas not being trained sufficiently now (e.g., close air support). **The technology permits coordinated training among the Services while individual Service elements remain at their home stations** [Simpson et al. 1995].

Command and Staff Training

Command and staff training occurs within constructive, live, and virtual simulations. The participating commanders and staffs range from the lowest to the highest echelon and from a single Service up through Multi-Service, Joint, and Combined commands. The most economical way to conduct such training is with constructive simulations, as they enable commanders and staffs to experiment without the cost of fuel, ammunition, and military personnel. Command and staff training does occur during live and virtual simulations, but usually these simulations are intended to train all participants at all levels. Because of their economy and relative ease of implementation, constructive simulations have proliferated in many different training domains. The *Catalog of Wargaming and Military Simulation Models (12th Edition)* [JCS 1993] lists 528 different models, all of which are currently in use.

Single-Service Training

- The 1990 REFORGER (Return of Forces to Germany) exercise made extensive use of constructive simulation to train leaders at brigade, division, and corps level. Benefits of such training were emphasis on battle planning, staff procedures, and command and control; more efficient use of training time; focus on higher echelons that would otherwise be cost prohibitive; reduced adverse environmental and political impacts. **The 1990 exercise saved more than $4M in transportation and cargo handling costs as compared to costs historically** [GAO 1991]. In 1992, constructive simulation was used to **avoid $34M in costs** as compared with the equivalent exercise done without simulation in 1988. Participants also believed that the training of staffs and planners involved was improved [Simpson 1995].
- The General Accounting Office noted that at the brigade level and above, simulations can be used to improve the decision making skills of senior battle officers before they command units in large scale training exercises [GAO 1993].
- Formal evaluations have demonstrated that constructive **simulations train commanders and staffs effectively and are relatively inexpensive**. The JANUS(A) is effective in training company level officers and platoon leaders on

current tactics and doctrine. The Brigade/Battalion Battle Simulation (BBS) has proven effective at training brigade and battalion staffs [Bryant et al. 1992].

Multi-Service and Joint Training

- The Defense Science Board concluded that **computer-based, simulated scenarios offer the *only* practical and affordable means to improve the training of Service operational commanders, their staffs, and the commanders and staffs who report to them**. Battle simulation offers the only opportunity to practice the use of certain weapon systems, sensors, tactics, and techniques against a skilled adversary [Defense Science Board 1988].
- Agile Provider (AP) was a Joint exercise sponsored by USACOM replaced by Unified Endeavor (UE) in 1995. AP was a field exercise last held in 1994. Unified Endeavor was supported by a Joint Training Confederation (JTC) of models interacting through the aggregate level simulation protocol (ALSP). The models replace ships at sea and flying hours, and focus on the primary training audience, the JTF commander and staff. **Total costs for AP-94 were $40M** with $8M in strategic lift costs. **UE-95's costs totaled $2.9M** with approximately $.5M in strategic lift. Approximately 85% of the UE-95 participants rated their training good and 82% rated it better than a similar field exercise like AP-94. The conclusion was **better training at 7.5% of the cost.**

ANALYSIS APPLICATIONS OF M&S

Analysis for Combat Operations

M&S contributes to innumerable decisions involving system evaluation and force sizing. In addition, it contributes significantly to combat operations. In 1990 and 1991, the Air Force Studies and Analyses Agency (AFSAA) performed a series of Gulf War analyses that Lieutenant General Glosson (then chief of CENTAF Special Projects) asserts "...**saved literally hundreds of lives**."

- A team of AFSAA analysts quickly deployed to the Air Force Operations Center in Riyadh, Saudi Arabia, where they analyzed the air campaign both before and after it began. For their analyses, they used primarily the Army's Space and Strategic Defense Command's EADSIM (also called the C3ISIM) model.
- At the time, EADSIM was a new model to AFSAA selected because it did an excellent job of analyzing command, control, and communications. It is a hybrid model with Monte Carlo and deterministic features. The combat operations planners were able to watch a preview of the attack as it unfolded in a way that graphically revealed the plan's strengths and weaknesses. Since the modeled air defenses, unlike the actual air defenses, acted in a rational manner, the simulation results showed a worst case scenario for the actual air assault.
- One main contribution was to choreograph the masses of aircraft into and out of the Kuwaiti Theater of Operations avoiding mid-air collisions and scheduling the rendezvous of tankers with attack aircraft. Planners also analyzed the best use of defense suppression assets, and alerted planners of missions that were too hazardous for some aircraft.
- For instance, their analyses indicated that it would be too dangerous to carry out plans to send A-6 and Tornado aircraft directly over Baghdad. As a result, only F-117 stealth fighters (none of which were lost), were assigned targets in that highly defended area. Undoubtedly, these changes saved lives and the needless loss of aircraft. When planners determined that SCUD sites in Western Iraq were too well defended and (as existing prior to the attack) too hazardous for F-15E attacks, defense suppression missions were reconfigured to correct the problem. When aircraft losses occurred, computer simulations were used to help determine the most likely cause so that later missions could be made less dangerous. To ensure that aerial tankers would make their rendezvous with fighters in need of refueling, missions were played out in advance. Attacks were carefully choreographed to avoid mid-air collisions, especially the first day's intense activity.

ABBREVIATIONS & ACRONYMS

AEDC	Arnold Engineering Development Center (US Air Force)
AFSAA	Air Force Studies and Analyses Agency
AP-94	Agile Provider 1994
ARI	Army Research Institute for the Behavioral and Social Sciences
BBS	Brigade/Battalion Battle Simulation
C4I	Command, Control, Communications, Computers and Intelligence
CCTT	Close Combat Tactical Trainer
CENTAF	Central Allied Forces
CGF	Computer Generated Forces
COEA	Cost and Operational Effectiveness Analysis
COFT	Conduct of Fire Trainer
DDR&E	Director, Defense Research and Engineering
DIS	Distributed Interactive Simulation
DISA	Defense Information Systems Agency
DMDC	Defense Manpower Data Center
DMSO	Defense Modeling and Simulation Office
DT	Developmental Test
DTSEE	Director, Test Systems Engineering and Evaluation
FDT&E	Force Development Test and Experimentation
FFRDC	Federally Funded Research and Development Center
GAO	General Accounting Office
HWIL	Hardware in the Loop
ICOFT	Institutional Conduct of Fire Trainer
IDA	Institute for Defense Analyses
IOT&E	Initial Operational Test and Evaluation
ISMT	Indoor Simulated Marksmanship Trainer
IST/UCF	Institute for Simulation Technology/University of Central Florida
JCS	Joint Chiefs of Staff
M&S	Modeling and Simulation
MACS	Multipurpose Arcade Combat Simulator
MCMSMO	Marine Corps Modeling and Simulation Management Office

MILES	Multiple Integrated Laser Engagement System
MOE	Measure of Effectiveness
MOP	Measure of Performance
MSBTF	Modeling and Simulation Benefits Task Force
MSWG	Modeling and Simulation Working Group
NAWCWD	Naval Air Warfare Center Weapons Division
OPTEMPO	Operating Tempo
OSD	Office of the Secretary of Defense
OT	Operational Test
OUSD (A&T)	Office of the Under Secretary of Defense for Acquisition & Technology
OUSD (P&R)	Office of the Under Secretary of Defense for Personnel & Readiness
P&L	Production and Logistics
PGTS	Precision Gunnery Training System
R&D	Research and Development
RDEC	Research, Development and Engineering Center (US Army)
REFORGER	Return of Forces to Germany
RFI	Request for Information
ROI	Return on Investment
SETS	Squad Engagement Training System
SIMNET	Simulator Networking
SMART	Simulation and Modeling Anchored by Real Testing
STX	Situational Training Exercise
T&E	Test and Evaluation
TECOM	Test and Evaluation Command (US Army)
TILV	Target Interaction, Lethality and Vulnerability
TF	Task Force
TOE	Tube launched, optically tracked, wire guided
UCOFT	Unit Conduct of Fire Trainer
UE-95	Unified Endeavor 1995
USACOM	United States Atlantic Command
USAMICOM	US Army Missile Command
VV&A	Verification, Validation, and Accreditation

VPG	Virtual Proving Ground

APPENDIX A

MEASURES OF EFFECTIVENESS

Problem Statement

The Defense Modeling and Simulation Office has articulated a need for assessing the impact of modeling and simulation (M&S) on the full range of Department activities. The impact of using M&S is difficult to quantify for several reasons: acceptable effectiveness metrics are lacking, supporting data are difficult to unearth, and in some cases it is not possible to identify a baseline from which to measure.

One member of the Task Force[1] examined 30 programs to determine how users of M&S measured their success. Those programs are listed and summarized after the exposition of findings. A second member of the Task Force presents a complementary discussion contained in Appendix B.

Findings from 30 Case Studies

Technical Areas, Functional Areas and DoD M&S Objectives

M&S capabilities fall into one of two broad areas. The Technical Area deals with mechanisms that make the M&S application work, while the Functional Area considers how the M&S application will be used. The two areas are discussed below.

- Technical Area: as mentioned above, this area deals with the inner workings of the M&S tool. The particular topics in this area are:

  Architecture: the high-level system and software design of the M&S tool

  Computer Generated Forces: the representation of constructive simulations

  Environmental Representation: how the real world is portrayed in the synthetic environment; the effects of weather, terrain, obscurants, and their interaction with the exercise entities

  Information/Database: methods for M&S tools to store or access information; data modeling

  Interoperability: how various M&S tools interface and operate together

  Networking: how information is shared among physically remote M&S tools

  VV&A: Verification, Validation & Accreditation, the process of giving M&S tools an official stamp of approval

  Instrumentation: details of the infrastructure needed to incorporate live entities into the synthetic environment; the hardware and software that allows real personnel and platforms to send their state variables between M&S tools, typically through electronic messages

[1] Mr. Matt Aylward of MITRE, Quantico.

- Functional Area: M&S tools are used in the following ways:

  Analysis: by using M&S tools to conduct experiments, useful information can be extracted

  Training: by using M&S tools to enhance military readiness through training, mission planning and mission rehearsal

  Acquisition (research & development): M&S tools allow virtual prototyping, enhance brainstorming, and expand the number of design options that can be considered

  Acquisition (test & evaluation): both Developmental Testing (DT) and Operational Testing (OT) can be accomplished with M&S tools. Breadboards and brassboards, combined with stimulation from M&S tools, enhance the quality of DT and OT results

  Acquisition (production & logistics): using M&S tools for design, manufacturing, process analysis, and support planning

The DoD Master Plan lists objectives that M&S tools should achieve. Some of these objectives include Framework, Environmental Representation, Systems Representation, Human Behavior Representation, and Developer and User Needs.

The Search for Metrics

To capture the benefits of any investment, there must be agreement upon objective standards to measure the performance of a particular investment. This appendix proposes one such set, containing both quantitative and qualitative metrics. The projects listed in this appendix were selected to demonstrate the breadth of M&S benefits enjoyed by the DoD. A discussion of the candidate quantitative and qualitative metrics follows.

Quantitative Metrics: Technical Area

- Architecture: Decisions about programming architecture have far reaching impact. Not only are current and future M&S tools affected, legacy systems will also feel some impact. The metric in this area, *Percent of Legacy Migration*, reflects this impact. An underlying Measure of Performance (MOP) to this metric might be the amount of effort (measured in man-years or dollars) required to migrate a legacy system to the proposed architecture.
- Computer Generated Forces (CGF): Modular Semi-Automated Forces (ModSAF) is the current DIS compatible software for generating entity-level personnel and vehicles; M&S tools featuring highly detailed constructive simulation of combat at the lowest level rely on this software. Wargames also use CGF, but usually at some higher level of aggregation. Following the same rationale developed for Architecture, the candidate metric is *Percentage of Software Reused*.

- Environmental Representation: The candidate metric, *Stimulations*, reflects the role of the environment in the real world. Military personnel constantly adjust their plans according to the current or future environment conditions, sorties are canceled or diverted due to obscurants over the primary target, or offensive operations are delayed because the main supply route is a sea of mud. This metric can be measured in many ways. For example, how many environmental effects can the M&S application portray? How many variables does the application use to describe the water column near the landing beach? So, within the metric of Stimulation, a number of MOPs can be developed for a given M&S application.
- Human Systems Interface: There are quantitatively rigorous techniques for measuring how well a M&S application approximates the real system; accordingly, the candidate metric is *User Acceptance*.
- Information/Database: Ready access to information, especially that information stored in databases, is critical to reducing the overhead of M&S tools. As future tools will rely on information archived in databases of various designs, analytical support of decisions involving database design should consider migrating legacy databases. Possible metrics include; *Level of Effort Required for Collaboration* and *Reuse*, measured in man-years and cost avoidance, perhaps arising from the reuse of an existing database.
- Networking: The links connecting the various sites involved in a Distributed Interactive Simulation (DIS) exercise are vital for exercise success. When choosing among various network designs, the analysis should identify the least expensive choice that ensures a successful DIS exercise. Therefore, possible metrics are *Cost per Megabit per second* and *Latency.*
- VV&A: If there is a suitable M&S application in existence, and this application has undergone a formal VV&A process, then a particular project has no reason to create a new application. Developing a rigorous VV&A procedure, particularly one that is not onerous to the owners of M&S tools, would enhance software reuse; accordingly, a candidate metric is *Cost Avoidance*, arising from such reuse.
- Instrumentation: When conducting instrumented exercises, the Services can completely forgo live fire or use the position reporting capability to have positive location information on each player. In either case, the risk of fratricide or erroneous fire mission approval by the Fire Support Coordination Center is greatly diminished. Consequently, a candidate metric is *Risk Reduction*.

Quantitative Metrics: Functional Area

- Analysis: Current evaluations of military operations rely heavily on anecdotal information; if analysts have a mechanism to capture what really happened, the resulting conclusions could have a much smaller confidence interval. An assessment of how various proposed M&S tools allow the analyst to better understand a process, and perhaps conduct the analysis from "ground truth," would be of obvious interest. One candidate metric for this assessment is *Net Utility*.

- Training: By far the functional area to benefit the most from M&S tools, it also has the most easily defined metrics. Examples include *Cost Savings*, *Cost Avoidance*, and *Risk Reduction*; for example, it's hard to have a Class 'A' mishap in an F/A-18 simulator.
- Acquisition (research and development): Again, possible metrics include *Cost Savings*, due to postponing metal bending, and *Number of Options Considered*. The second metric for research and development (R&D) and production and logistics (P&L) captures the same idea; study a wide range of options for the proposed system in virtual reality before making the first bend in metal.
- Acquisition (test and evaluation): A number of case studies show that using M&S tools to prepare for DT and OT is advantageous. M&S allows for better designed tests, aids in training the test force, and identifies areas of deficient data collection. In some cases, such as testing software upgrades for the F-14, M&S is the only way to conduct DT due to the risks involved. Use of *Cost Avoidance* as a metric is well supported by these case studies.
- Acquisition (production and logistics): Possible metrics include *Cost Savings* and *Number of Options Considered.* The former accrues when production lines and consumption rates can be simulated, allowing problem identification and correction before the system is fielded. The latter is appropriate when application of M&S allows planners to explore a wide range of production options in detail, at a minimal cost, when compared to exploring the options with real world equipment.

Quantitative Metrics: DoD M&S Master Plan Objectives

The five criteria articulated under the M&S Objectives identified in the M&S Master Plan are closely related to the preceding sections on Technical and Functional Areas. Accordingly, a separate discussion of each M&S application area in unnecessary.

Qualitative Metrics: Technical Area

- Computer Generated Forces: The more realistic the portrayal by the CGF, the better the training; therefore, a candidate metric is *Training Quality*, as evaluated by the training audience.
- Environmental Representation: One possible metric is *Immersion*, a measure of how "real" the environment feels to the trainee.
- Human Systems Interface: How well the M&S application matches the feel of the simulated system is critical to positive training transfer. For this reason a candidate metric is *Ease of Use*, as evaluated by the trainee or instructor.

- Interoperability: Ideally, all Service models would interoperate readily, thus ensuring that the conclusions of a study sponsored by one Service would be acceptable to all Services. For example, the Office of the Secretary of Defense (OSD) can choose among a number of combat simulations to evaluate Service roles and missions; the selected simulation should neither over nor understate the relative value of any one Service. A candidate metric, *Level Playing Field*, embodies this idea of a neutral evaluation tool.

- Networking: If the M&S application is distributed, the network can either degrade or enhance the sensory experience of the participants. The quality of sensory stimulation, the feeling of *Immersion* experience by players, is a possible metric.

- VV&A: The proposed metric is *Enhanced Decision Support*. Here, the implicit assumption that a model that has completed a formal VV&A process produces results that are "more valid" than results from a non-VV&A process is extended one step further. "More valid" results will lead to better quality decisions.

- Instrumentation: A possible metric is *Merger of C4I with M&S*. For a variety of reasons, it is desirable for these two disciplines to merge, with M&S tools executing over current C4I systems. This metric estimates how closely a particular M&S application approaches this goal.

Qualitative Metrics: Functional Area:

- Analysis: As discussed in the quantitative section, much analysis of military operations is based on anecdotal data. Merging live, virtual, and constructive simulations with C4I systems would give analysts much better quality data with which to work. For these reasons the metric, *Data Quality*, is offered.

- Training: The candidate metrics are *Readiness* and *Unique Training*. Readiness is enhanced because the lower cost of training with M&S tools increases the quantity and quality of training opportunities. Unique training is possible with M&S because only electrons are in danger of getting killed. This latter metric speaks to the ability of M&S to present trainees with situations not seen in the real world outside of combat.

- Acquisition (research & development): The candidate metrics are *Brainstorming* and *Unique Capability*. Brainstorming refers to the ability of M&S to let program managers explore a much wider array of options before settling on one approach. In other cases the unique capability of M&S can't be reproduced in the real world.

- Acquisition (test & evaluation): Some possible metrics include *Developmental Test Planning*, *Operational Test Planning*, *Development of Measures of Performance (MOP)*, and *Development of Measures of Effectiveness (MOE)*. These metrics assess how well an M&S tool assists the T&E community in all of the foregoing tasks.

Qualitative Metrics: DoD M&S Master Plan Objectives

Just as the quantitative metrics for the Technical and Function Areas were applicable to the area of M&S Objectives, so too are the qualitative metrics.

Conclusions

In order to measure the impact of the many M&S applications in the DoD, we must first state our objectives in quantifiable terms. Only then can we assess our progress toward reaching those objectives. The metrics described here are offered as a strawman, sure to draw criticism, that will move us toward measuring the impact of M&S.

Projects Examined for Measures of Effectiveness

1. Indoor Simulated Marksmanship Trainer (ISMT): A stand-alone M&S application for training Marines in all aspects of small arms fire, as well as training small units in tactical engagements. A MITRE analysis indicated that use of the ISMT for a portion of annual weapons requalification could save significant amounts of money as a result of ammunition offset.

2. Deployable Forward Observer-Modular Universal Laser Equipment (DFO-MULE): A stand-alone M&S application, it is a training device for forward observers (artillery and mortar) and forward air controllers. It complies with current DIS standards. The DFO-MULE is being used in the Multi-Service Distributed Testbed. A MITRE analysis indicated that use of the DFO-MULE for required forward observer training could save significant amounts of money as a result of ammunition offset. Assuming a 10% offset in live fire tasks, savings in ammunition expenditures could recover the acquisition costs before the end of the second year [Fish 1995c].

3. Emerald Light: A Marine Corps proof-of-concept demonstration, it will instrument a training range at the Marine Corps Air Ground Combat Center (MCAGCC). Ultimately, the MCAGCC and the National Training Center (NTC) at Fort Irwin, CA, will be linked. This will allow the conduct of Joint exercises at both sites. During the exercise, participants will share a synthetic battlespace over the Defense Simulation Internet (DSI).

4. Synthetic Theater of War-Europe (STOW-E): An ARPA-funded effort that allowed the Army to link live, virtual and constructive simulations in order to conduct a large scale training event embedded within ATLANTIC RESOLVE. Comparable in size to REFORGER (Return of Forces to Germany), ATLANTIC RESOLVE demonstrated savings on the order of $37 million.

5. LeatherNet: An ARPA-funded project developed in concert with STOW-97, this project seeks to create a credible Marine Corps CGF at the level of the individual rifleman.

6. Turret Layout: Not really a project, as much as a study; this effort compared the use of M&S to prototype construction for developing modifications to the Abrams tank. The M&S application allowed more options to be considered, at a lower cost, in less time, while involving the user community; the benefits were clear cut and convincing.
7. Advanced Medium Range Air to Air Missile (AMRAAM): During the development of this system, investments in M&S tools allowed engineers to fly a complete mission profile in a virtual environment. The amount and quality of data made available from this investment far exceeded the telemetry from a real flight. Also, given the cost of each flight, and the number of flights required, actual flight testing was prohibitively expensive. Only M&S could satisfy the need for performance data at an acceptable cost.
8. AIM 9X: See discussion of the AMRAAM (*7*).
9. F-14 Software Test: The F-14, like most modern aircraft, relies on computers to execute its mission. Changes to its software are made continuously, yet each change could potentially result in non-desirable flight performance, such as crashes. For this reason, each change in software must be rigorously tested before the plane is flown. This testing must be conducted on the ground, with all flight control systems receiving accurate input stimulation. Only M&S tools can provide this input. In the absence of simulation, upgrades to the F-14 software would not be possible.
10. Advanced Amphibious Assault Vehicle (AAAV): While this program is moving forward under the conventional acquisition paradigm, it is concurrently looking at ways to use M&S tools to change the process. One example is the participation of the two automotive test rigs participating in the Virtual Proving Ground. In this project, data collected from real vehicles on the test track at Aberdeen Proving Grounds is compared to similar data produced by computer simulations. Successful completion of the Virtual Proving Ground project will give designers the ability to consider many design options without bending any metal.
11. Boeing 777: During the development of this aircraft, the Boeing Corporation re-invented itself: this is the world’s first paper-less airplane. Boeing’s corporate information system architecture allowed for extensive use of CAD, CAE, and CAM (computer-aided design, engineering, and manufacturing). Production of the aircraft was greatly simplified, as only a minuscule amount of time was lost to poorly fitting parts. In turn, this reduced the costs of production by reducing labor overtime charges and scrap rework.

12. New SSN Prototype: Employing simulation (see Boeing 777, (*11*)) for the follow on to the SSN-21 (Sea Wolf class), General Dynamics has already experienced cost avoidance on the order of tens of millions of dollars. Proposed changes to the weapon systems consoles, sensor suites, or engineering plant can be completely explored in a virtual environment before any metal is bent. Compared to previous construction methods, large costs are avoided. This is an example of how the corporate information system architecture can have a significant impact on the bottom line.

13. B-2 Mockup: In a vein similar to Boeing and General Dynamics, Northrop designed its CISA to gather information throughout the company. As a result, the mockup of the B-2 was so close to the design resident in the CISA that future mockups may be eliminated entirely.

14. SEEK IGLOO: An Air Force project to deploy warning radars, the concept of employment called for manned installations. A MITRE simulation determined the radars were much more reliable that assumed. This led to a different concept of employment, unmanned radar installations of smaller size. Large savings from cost avoidance were realized.

15. F-16 Operational Test Scenario Development: By using simulators, the OT project officer was able to realize a number of benefits. First, the test team was fully trained in the scenario for the OT, increasing the efficiency of the test. Second, the test crews were able to show the project officer what performance measures were truly important, leading to a modification of the test scenario. In this way, simulation led to a higher quality OT.

16. Forward Area Air Defense System/Air Defense, Anti-Tank System Measure of Performance (FAADS/ADATS): Since neither of these systems exist, the project officer was stymied in attempts to develop appropriate measures of performance (MOP). The use of simulation allowed the project officer to clarify the concept of employment and develop worthwhile MOPs.

17. Non Line of Sight OT (NLOS OT): The NLOS is new type of anti-tank weapon that allows for precision attack of armored vehicles by a gunner in full defilade. Facing problems similar to the FAADS/ADATS, the project officer turned to simulation. Again, the existence of sophisticated M&S tools led to high quality OT of a future system.

18. Virtual Proving Ground: This is an effort between the Army’s Aberdeen Proving Grounds and the University of Iowa. It is an attempt to create a synthetic environment for testing vehicles. The goal is to allow engineers to fully explore system design (e.g., the HMMWV) before any metal is bent.

19. Joint Warfare Concept Analysis-Operations Research (JWCA-OR): An effort to improve the quality of Joint analysis. For this work, it is essential that all Services are represented, as JWCA-OR supports force structure decisions, aids in developing Joint doctrine, and guides force allocation to the CINCs. Currently, Service representation largely involves legacy systems.

20. Joint Warrior Interoperability Demonstration (JWID): A series of demonstrations sponsored by the Chairman, Joint Chiefs of Staff. They are primarily concerned with C4I; future demonstrations will see the integration of M&S with C4I.

21. Standard Interchange Format (SIF): Developed by the Institute for Simulation and Training (IST), SIF allows existing databases to interface with M&S tools. The use of SIF generates savings for each project by avoiding a quantity of many years normally required to develop a custom interface.

22. B-52 Data Study: Undertaken by the Air Combat Command (ACC) during Operation Desert Shield/Desert Storm, this effort collected a wide array of operational data from the bomber force. This data is potentially very useful for a number of different M&S functional uses.

23. Jedi Knights: A colloquialism that refers to a group of Army officers who provided operations research support to CGCENTCOM in theater. This group was drawn from the Command and General Staff College and were experts in the wargame TACWAR. Prior to the start of hostilities, the Jedi Knights played TACWAR manually and compared the results to the computer results from the same scenario. They judged the TACWAR output credible and proceeded to use TACWAR for operational support. The Jedi Knights are an example of the effort and benefits associated with VV&A of simulations.

24. Desert Storm Operations Research: The CGCENTCOM, as well as subordinate commanders, made extensive use of operations research personnel as plans for Operation Desert Storm were developed or executed. Analytic support was provided from the United States, as well as from operations research (OR) cells in theater. Plans were developed, analyzed and modified in a greatly truncated cycle. Without sophisticated M&S tools, the OR cells would have been unable to respond to the needs of the operational commanders. The benefit of simulation was especially evident in planning and conducting the air campaign.

25. Joint Surveillance Target Acquisition Radar Terminal Emulation (JSTAR TE): Originally intended as an adjunct to the JSTAR program, the JSTAR TE allowed the JSTAR to reach operational capability in time for Desert Storm, six years ahead of schedule. This was a great success, both for the war effort as well as for the program.

26. Defense Information Systems Agency (DISA): DISA is responsible for developing the communication links necessary to connect the far flung activities of the DoD. In this task, DISA has used a number of M&S tools to consider various alternative methods of linking the DoD nodes. There are several documented instances of cost avoidance that are the direct result of using M&S.

27. Simulation Utility Management System (SUMS): SUMS is an Air Force effort to develop an M&S tool to assess the effects of changing manpower policies and programs. It also allows personnel planners to consider various scenarios regarding the nature of the civilian labor pool.

28. Virtual Medicine: This project is still in the basic research phase, but if offers tantalizing benefits. Battlefield surgeons could operate without subjecting the wounded to the trauma of transportation to a field hospital. This multiplies the effectiveness of each surgeon, while reducing demands on the transportation system and eliminating a lucrative rear area target, the large field hospital.

APPENDIX B

METHODS OF ANALYSIS

One member of the Task Force gathered examples of data and cost-effectiveness analysis with which she was familiar.[2] Through three case studies, she shows alternative methods of calculating cost effectiveness using the same data. Subsequently, she generalizes her observations and discusses a range of effectiveness measures and cost components.

Alternative Calculations for Three Case Studies

The separate case studies were chosen because of the availability of data to the author and their ability to demonstrate the various methods of analysis. The data come from, or have been examined by, the relevant personnel. Four different methods of calculating cost effectiveness are shown, each providing a different result. The first, and most common, method calculates cost savings or avoidance, and is usually based on the assumption that live and simulated events are completely interchangeable. The second method is break-even analysis that determines how many live events must be replaced by simulated events to recover capital investment and operating costs in a given period of time. The third method is based on the assumption that finding errors early in the acquisition process are less costly to repair than finding them later in the process. The final method of comparing alternative events is to compare their costs and their effectiveness separately. While the most general method, it is difficult to implement due to the inability to adequately measure the effectiveness of an event. This final method would allow comparison of an alternative, M&S-supported event with a baseline event where the alternative was more costly but provided better training for example. Military experts could then decide whether the training increment (decrement) was worth the additional (lower) cost.

AMRAAM Hardware in the Loop

The AMRAAM Hardware-in-the-Loop (HIL) facility at Point Mugu is employed in the ongoing evaluation of missile guidance and control system performance. The facility includes a flight simulator table, anechoic chamber, target simulators, special interface hardware, and an instrumented missile. The facility can be used for additional applications, but only its use for testing the AMRAAM is considered here. Its primary cost components are shown in Table B-7.

Table B-7: AMRAAM HIL Costs

BRAC Replacement Cost[3]	$23.7 M
Yearly Operating Cost	$930K ($1M)

[2] Ms. Michelle Bailey of NAWCWD, China Lake, CA.

[3] Base Realignment and Closure (BRAC) replacement costs represent the cost to replace a facility and not the cost of original development and maintenance. This metric is used here because it is a certified figure with a definite meaning applied uniformly across the country.

Number of Tests Run/Year	8,400

Using the assumption that all firings are live, shown in Table B-8, we calculate an extremely favorable cost savings. While 8,400 *simulated* firings per year are possible, a program simply could not afford 8,400 *live* firings. Even so, several M&S cost savings we have gathered have been calculated using the type of assumption.

Table B-8: AMRAAM Example Savings Calculations

Cost Savings Method (assumes all live firings)	
Cost per firing	$40K
Cost of missile	$250K
Total cost per firing	$290K
Number of tests	× 8,400
Total savings	$2,436M
Could a program do 8,400 firings, let alone in one year?	

For this case, break-even analysis may be more meaningful. An example is shown in Table B-9. In this case recovery of BRAC replacement costs occurs in 10 years, assuming that 12 firings per year are simulated, a far more reasonable assumption. Further, assuming 3 or 4 firings per year beyond the 10-year break-even point will recoup facility operations and maintenance costs. The conclusion, then, is that 12 firings per year will recoup the capital investment in 10 years, and the other 8,388 simulated firings are value added, i.e., contribute more effectiveness. It must be remembered that an AMRAAM missile costs considerably more during its development and early production. Consolidation of the earlier missile costs with later production costs would shorten the payback time.

Table B-9: AMRAAM Example Break-Even Calculations

Time to Break-Even Method
• Number of firings required for break-even in 10 years at 12 firings per year (23.7 + 10 × (operating costs) = $34M / $290K = 117 firings in 10 years) • Number of firings saved per year to maintain cost effectiveness ($1M / $290K = 3.45 firings per year)

A third alternative method of calculating the cost effectiveness of this type of facility is to record the number of errors found during HIL testing that would have caused a firing failure. Live firings are an expensive way of finding errors. We do not have the data required to conduct this type of analysis. However, we can approach it using the F/A-18 WSSF data.

F/A-18 Weapons Software Support Facility (WSSF)

The F/A-18 Weapons Software Support Facility (WSSF) at China Lake, California, is used for integration, checkout, and V&V of avionics software with actual avionics hardware operating as a total aircraft system. The WSSF is actually several facilities containing avionics hardware, simulations of flight dynamics, weapons simulations, and operator consoles. Table B-10 shows the WSSF cost factors used in the following example calculations.

Table B-10: F/A-18 WSSF Cost Factors

BRAC Replacement Cost	$54 M
Yearly Operating Cost	$6 to 8 M
Number of Test Hours per Year	over 6,000
Lab Costs per hour (F/A-18)	$930
Lab Costs per hour (other aircraft)	$1550
Ground Costs per Hour	$100
Flight Costs per Hour	$2,800

The WSSF is also used by weapons programs for integration and checkout of their aircraft interfaces. It has also been used to supply simulated aircraft for other tests. The cost savings we are going to compute are just the savings for the F/A-18, not all programs which use the facility. It is important to note that the cost per flight used here is the actual figure charged to the project. The true fully amortized cost of keeping an F/A-18 in the air and flight ready is probably much higher. The topic here is methodology.

Table B-11: F/A-18 WSSF A/B Software Upgrade

Cost Component	Hours Expended	% Errors Found
Lab Hours	1,084	73%
Ground Hours	81	2%
Flight Hours	195	11%
Other Methods[4]	?	14%

[4] Includes code reviews and other paper-based checks.

Table B-12: F/A-18 WSSF C/D Software Upgrade

Cost Component	Hours Expended	% Errors Found
Lab Hours	4,957	61%
Ground Hours	440	4%
Flight Hours	966	13%
Other Methods	?	22%

The more errors that can be found in the early stages of development using WSSF, the cheaper the overall program without even considering safety issues. Ground tests are relatively cheap, but they can only be used for simple power checks. They are included here so that total errors add up to 100%.

Table B-13: WSSF Example Savings Calculations

Cost Savings Method (assumes use of flight hours for all lab debug)		
Cost of flight hours (A/B)	1,084 hrs × \$2,800/hr	\$3M
Cost of flight hours (C/D)	4,957 hrs × \$2,800/hr	+ \$14M
Total cost of flight hours		\$17M
Annual operating costs		– 7M
Savings per year		\$10M
However, there are not enough local planes to fly 6,000 hours in one year.		

The break-even viewpoint, based on replacement and operating costs, yields a more reasonable number of flight hours, but even that is difficult for one test facility to bear. If each lab hour equated to a flight hour, we would need more than one facility testing the software, or we would need more F/A-18s dedicated to software integration and test. The WSSF actually has several labs which may be used in parallel.

Table B-14: WSSF Example Break-Even Calculations

Time to Break-Even Method	
Replacement costs	\$54M
Maintenance costs (\$7M per year)	\$70M
Total costs	\$124M
Cost per flight	\$2,880
Have to save 4,400 flights per year for 10 years (\$124M/2,880 = 4,428)	

The real value added of the WSSF is that an aircraft as complex as the F/A-18 is not possible without this type of test facility. We could not fly enough to test

it. There is a danger in just looking at cost savings as the measure of whether or not we invest in M&S. As we demand more from our warfighting systems—safer, more accurate, more environmentally friendly, more stealthy, longer range, etc.—we will have to demand more from our test and training systems.

Kernel Blitz

Kernel Blitz was a fleet training exercise (FLEETEX) including live ships, submarines, aircraft, and land troops. The simulation portion augmented the fleet with additional synthetic ships, submarines, aircraft, and weapons. The simulation center used several existing computer facilities (including both coasts) and existing communications capability to link to platforms. A purpose of the exercise was to show that the use of simulated assets could add realism and complexity to training exercises.

Table B-15: Kernel Blitz

Simulated Assets	
Ships and submarines	
(23 platforms × 2 days × $100K/day)	$4.6M
Aircraft	
(27 platforms × 4 hrs × $3K/hr)	+ $0.3M
Weapons	
(23 weapons × $500K/weapon)	+$11.5M
Costs	
BFTT Enhancements	$350K
Total Savings	$16M

The Battle Force Tactical Trainer (BFTT) existed prior to Kernel Blitz but was enhanced for this exercise. During Kernel Blitz, 33 real ships and submarines were used. The simulated assets significantly increased that number. The fleet commands will have to answer the question of whether they would ever put 55 ships and subs into a training exercise. The commanders at sea quickly forgot who was real and who wasn't.

The $500K per weapon might seem high to some, but it is the value used by the BFTT office (the AMRAAM is running about $250K per copy).[5] Regardless of what we think about the full cost of all the simulated assets, the $350K modification costs are impressively low (they represent only those costs charged to Kernel Blitz via BFTT). If we counted only the use of two ships for two days ($400K) the Navy would recoup its investment.

[5] There may likely be different cost factors included in the $500K and $250K values

An analysis by the Center for Naval Analyses[6] states that "**At this point, simulation should be viewed as enriching training and increasing readiness rather than reducing costs.**" The CNA analysis also specified much greater costs. However, the purpose of this discussion is not to determine the cost effectiveness of Kernel Blitz, but to demonstrate the alternative analytic methods.

Can or Should Cost Savings be the Ultimate Measure of M&S Impact?

There are four basic categories of effectiveness measures obtained from applying modeling and simulation—**doing it better**, **doing it faster**, **doing it cheaper**, **doing it at all**. By "doing it better," we mean that the quality of the product or the quality of the processes employed is improved through the application of M&S. This is sometimes hard to measure in terms of dollar savings. **What value do we put on safer processes? We know how to determine the cost of a disaster, but what about the cost of near misses?** Is there a savings from reducing the number of near misses? We can usually obtain dollar savings for more accurate testing, earlier discovery of problems, and repeatability of testing, although there will be some subjectivity in the figures.

Simulations make it possible to conduct training events or test events that would not be possible or affordable if conducted live. **Is it reasonable to compare the cost of a simulated event to the cost of a live event that never would have occurred?** Better would be to compare the costs and effects of two realistic but different live and simulated events. However, we often lack the appropriate effectiveness measures.

Sometimes, it just wouldn't be possible to conduct a specific test, or train for a specific situation, without simulation. For instance, in testing seekers, there are neither enough test points nor space to hook up test equipment to obtain all the information needed about the behavior of the hardware. By using a simulation, we have access to all parameters. For aircraft, we want them to be able to withstand a certain amount of G-loading, but **to actually test that would mean risking the loss of an aircraft at the edge of its envelope**—so we simulate the effects of Gs through application of M&S.

When classifying the effectiveness of M&S, much depends upon its application. Is it a wargame or an engineering simulation? Identifying effectiveness measures for the different M&S applications would make it easier for M&S users to keep track of the impact. For instance, wargames may save money by identifying shortfalls of existing weapons by pointing out tactical solutions vice acquisition solutions, or by identifying a set of equally effective solutions from which the least expensive can be chosen. Engineering simulations save money by enabling faster design. **But what is the value added by increasing the number of alternatives considered through simulation?**

---

[6] Thomas K. Neuberger and Dennis P. Shea, *Applying Synthetic Environments to Operational Training: A Perspective from Kernel Blitz 95*, CNA CAB 95-58, September 1995.

Candidate Measures of Effectiveness

Table B-16 lists candidate measures of effectiveness (MOEs) and how they might be derived or calculated. The application space is mission planning.

Table B-16: Mission Planning

	Measure of Effectiveness (MOE)	**Determination of MOE**
Faster	Actual time savings of commanders doing the strategy, per mission	Review mission planning times
	Value added of “quick reaction” capability: shorten war, avoid casualties	Wargame with and without M&S capability
Better	Value added of additional (on-line) information to mission planners	Review costs of mistakes, wargaming
	Value added of considering additional strategies	Wargame with and without M&S capability
	Value added of considering multiple enemy reactions to strategy	Review costs of unexpected reactions
Cheaper	Cost savings of using new methods	Review cost of current equipment and compare to projected costs
At All	Value added of retargeting mission en route	Number of “wasted” missions
	Value added of more detailed planning	Operator assessment of mistakes caused by ambiguity
	Value added of automatic recording/ review of strategies, scenarios and lessons learned	Evaluate mission planning training methods

Table B-17 lists candidate MOEs for M&S tools that support analysis in support of Research, Development, Test and Evaluation (RDT&E). Examples of this type of analysis includes cost estimation, technical effectiveness evaluation, and cost and operational effectiveness analysis (COEA).

Table B-17: RDT&E, Analysis

	Measure of Effectiveness (MOE)	Determination of MOE
Faster	Better adherence to schedule	Review daily program expenditures
	Use of virtual prototyping	Look at turn-around time for physical models
Better	Value of adding more detailed analysis	Review number of design, software, planning changes
	Value added of considering more alternatives	Review pre-planned product improvement, cost reduction, packaging efforts
	Value added of making better decisions	Estimate of unknown unknowns, number of backup plans used for risk mitigation
Cheaper	Cost savings of using new methods	Review cost of current equipment purchase/use/maintenance and compare to projected costs
At All	Value added of "executable requirements"	Costs of erroneous requirements: suits, redesigns, ambiguities
	Value added of operators of virtual prototyping	Costs of failing OPEVAL

Table B-18 lists candidate MOEs for M&S tools employed in the design phase of RDT&E. Examples include trade-off studies, engineering simulations, parametric optimization, maintenance planning, logistics planning, and production planning.

Table B-18: RDT&E, Design Phase

	Measure of Effectiveness (MOE)	Determination of MOE
Faster	Reduction of design iterations	Compare to similar efforts
	Automatic design documentation	Compare to manual methods
Better	Incorporation of maintenance, logistics, and production considerations	Estimates of reduced life cycle costs from what simulations pointed out
	Incorporation of maintenance, logistics, and production considerations	Estimates of reduced life cycle costs from what simulations pointed out
Cheaper	Use of virtual prototyping	Cost of physical models
At All	Evaluation of designs under more situations	Estimated costs of design failure under those situations

Table B-19 offers seven MOEs for M&S tools used in test and evaluation.

Table B-19: Test and Evaluation

	Measure of Effectiveness (MOE)	Determination of MOE
Faster	Better adherence to schedule	Daily cost of ranges, program slips
	Better use of flight test time	Percent of test time wasted
Better	Value added of “monte carloing” test conditions	Percent of operational requirements not physically tested, but inferred from testing
	Value added of rehearsing test	Percent of tests wasted
	Earlier identification of problems	Look at cost/spending curves for phase of project, look at cost of ECPs by phase
Cheaper	Use of virtual prototyping	Cost of physical models
At All	Evaluation of designs under more situations	Estimated costs of design failure under those situations

Table B-20 lists seven candidate MOEs for M&S application to training and how those MOEs might be determined.

Table B-20: Training

	Measure of Effectiveness (MOE)	Determination of MOE
Faster	Cost savings for fewer training days	Review average number of days for specific training
	Value added of training en route	Percent delay in deployment due to training
Better	Value added of exposing trainees to more situations	Review of operator errors
	Total assessment of trainee progress	Evaluation of individualized training to graduate some individuals early
Cheaper	Cost savings of using new methods	Review cost of current methods and equipment
At All	Individual remedial training	Review number of "flunked" trainees
	Virtual reality training in hazardous situations	Review casualties, accidents due to operator error

And finally, Table B-21 lists candidate MOEs for M&S tools in support of military operations.

Table B-21: Support to Military Operations

	Measure of Effectiveness (MOE)	Determination of MOE
Faster	Logistics routing	Time saved with better method
Better	Weapons mix studies, both platform and individual	Enhancement of capabilities from tailored mixes
Cheaper	Reduction of personnel required to do analysis	Amount of analysis based upon simulation
At All	Decision aids	Benefits of faster, better decisions

Identification of Expenditures

The previous sections identified several measures of effectiveness for M&S. This section takes a closer look at the costs. Ideally, cost estimation would be the responsibility of each program manager when determining whether to pursue M&S versus other options. Too often we examine only the costs of building the simulation (or enhancing an existing one) and forget about the cost of V&V, training, operation, and maintenance.

An additional shortcoming in assessing the cost effectiveness of M&S is that the cost of failure is rarely captured. With the large up-front costs associated with some M&S tools, a considerable amount of money can be spent before determining that the tool simply won't work. Unfortunately, these lessons learned rarely get publicized, so we "learn" the same lessons repeatedly.

The costs of an M&S tool are less dependent upon its application domain and more dependent upon its physical implementation. If a simulation is entirely software, its costs can be identified in the same fashion as any other software system. The same is true of hardware-in-the-loop and live simulations.

The greatest difficulty in acquiring data is getting the right data and understanding its meaning and limits. It is imperative that the M&S community decide what data it needs and provide guidelines to program managers on how to record that data. It doesn't have to be difficult if people know from the beginning what is needed.

The second problem is making sure the data are used correctly. Good data used against the provider will not engender more good data. This is a political problem and hence more difficult to solve than the first.

Table B-22 summarizes the identification of expenditures.

Table B-22: Identification of Expenditures

	Area of Cost Saving/Avoidance	Determination of MOE
Build	The simulation is a product, just like a weapon system	Treat like an acquisition program, important to do a feasibility study
	Costs of "productizing" M&S; making it usable by several people	May be applicable if using a legacy simulation
	Integration of simulation with other simulations or hardware	May have to pay to modify interface software or equipment
Prove	VV&A	Probably 50% of acquisition costs, simulations are software intensive
	Operator acceptance—do the users/ customers believe the results?	Need operator involvement from start, increasing acquisition costs
	Validation testing	Actual hardware tests to validate the models may include live ordnance firings—number and type need to be determined during program planning
Use	Training of users	Recurring cost—users will change
	Training of facilitators—people who train the users & run the simulation	Recurring cost, but at a slower pace than training of users
	Computer time	Lease or computer costs
	Equipment storage, access to space	May be leased
	Scheduling time	Delay to program because simulation facility was not available exactly when needed
	Setup costs	Simulation may be scenario dependent or user tailored
	Duplication of equipment	Users may choose to purchase their own systems—they also need to duplicate facilitator costs and maintenance costs
Feed	Population of databases	Dependent upon scenario, access/cost of data
	Equipment maintenance contracts	Recurring costs
	Configuration management	Recurring costs
	Depreciation of equipment	Sometimes applicable
	Lease of communications lines	Recurring costs
	Update of databases	Scenario dependent
	Software support activity	Make modifications, upgrade, fix
	Revalidation	Each time a change is made
	Maintenance of libraries	Baselines and distributing releases

	POC for questions	Necessary if system at multiple sites